CLAIMS

What is claimed is:

1   1.   A system for mining data comprising:

2   a data store including data having a number of items;

3   a mining application to mine data in the data store, the mining application

4 including logic, the logic, when executed, is to:

5   identify a number of frequent items of the data store;

6   compute a probe structure based on the number of identified frequent

7 items; and,

8   partition the data according to content of the probe structure;

9   wherein the mining application uses the probe structure to build a frequent

10 pattern tree (FP-tree); and

11   a memory for storing the probe structure and the FP-tree.

1   2.   The system of claim 1, wherein the data of the data store includes a number of

2 transactions, wherein each transaction comprises a unique sequence of items

3 identified by the logic when identifying the frequent items of the data store.

1   3.   The system of claim 2, wherein the logic is to partition the transactions

2 according to content of the identified frequent items to obtain the probe structure,

3 wherein the probe structure includes combinations of the identified frequent items and

4 the number of occurrences of one or more content-based transactions.

1   4.   The system of claim 3, wherein the logic orders the identified frequent items

2 based on an occurrence frequency of each identified item in the data store.

1   5.   The system of claim 3, further comprising a heuristic algorithm, wherein the

2 heuristic algorithm is to group the one or more content-based transactions into

3 approximately equal groups.

1   6.   The system of claim 1, further comprising a master processor and one or more

2   slave processors, wherein the master processor is to distribute a group of transactions

3   to the one or more slave processors to build the FP-tree.

1   7.   The system of claim 6, wherein the one or more slave processors build a part

2   of the FP-tree based on the grouping of content-based transactions.

1   8.   The system of claim 7, wherein the multiple processors mine the FP-tree to

2   determine unique information about the items of the data store.

1   9.   The system of claim 1, further comprising a multi-core system architecture.

1   10.   A system for mining data, the system comprising:

2   a database including a number of transactions;

3   at least one processor to perform mining operations on the database, the at

4   least one processor is to execute content-based partitioning logic on the transactions,

5   wherein the content-based partitioning logic is to partition the transactions according

6   to content based on a number of identified frequent items to obtain a probe structure;

7   and

8   a memory to store the probe structure.

1   11.   The system of claim 10, the probe structure further comprising a probe tree

2   and probe table, wherein the probe tree and probe table further comprise 2^M branches,

3   wherein M corresponds to the number of identified frequent items.

1   12.   The system of claim 11, wherein the memory further comprises shared

2   memory to store the probe tree and probe table.

1   13.   The system of claim 11 further comprising multiple processors to recursively

2   mine the database, wherein each processor shares a substantially equal load based on

3   a grouping and distribution of the 2^M branches.

1   14.   The system of claim 13, the multiple processors further comprising a master

2   processor and at least one slave processor to perform mining operations, wherein the

3   master processor distributes operations to the at least one slave processor when

4   building a frequent pattern tree (FP-tree) using the probe structure.

1   15.   A method for mining data of a database, comprising:

2   identifying frequent items of the database;

3   building a probe structure based on the identified frequent items, wherein each

4   branch of the probe structure includes a number of identified frequent items based on

5   content;

6   grouping the branches of the probe structure based on the content of each

7   branch; and

8   building a frequent pattern tree (FP-tree) from the probe structure.

1   16.   The method of claim 15, further comprising scanning a first portion of the

2   database when identifying frequent items of the database, and scanning a second

3   portion of the database when building the probe structure, wherein the probe structure

4   includes an associated number of counts with each branch of the probe structure after

5   scanning the second portion of the database.

1   17.   The method of claim 15, further comprising building the probe structure to

2   include a probe tree and probe table, and using the probe tree and probe table to build

3   the FP-tree for mining the FP-tree to determine frequent data patterns.

1   18.   The method of claim 15, further comprising distributing each group of

2   branches to an associated processor before building the FP-tree.

1   19.   The method of claim 18, further comprising using a master processor to

2   distribute each group of branches to one or more slave processors, and using the one

3   or more slave processors to build the FP-tree.

1   20.   The method of claim 15, further comprising partitioning the database

2   according to content of the identified frequent items to obtain the probe structure,

3   wherein the probe structure includes combinations of the identified frequent items and

4   the number of occurrences of one or more content-based transactions.